Terrestrial mass extinctions and galactic plane crossings

IN their attempt to build a case for a "galactic plane crossing" explanation for terrestrial extinction, Rampino and Stothers[1] present a statistical argument that is seriously misleading. The crux of the empirical component of their argument is a claimed agreement between two columns of dates (their Table 1):

Table 1 Dates from ref. 1

Mass extinctions (Myr BP)	Galactic plane crossings (Myr BP)
11	0
37	31
66	64
91	100
144	135
176	166
193	197
217	227
245	259

They begin by calculating the correlation coefficient of these two columns to be $r=0.996$. It is a mistake to attribute any meaning to this number: the correlation of any two monotonically increasing sequences is bound to be high. This reflects only the high serial correlation of the separate series, and in no way indicates any connection between the two series other than a common monotonicity. For example, the correlation between the list of mass extinction dates and the first nine prime numbers (2, 3, 5, 7, 11, 13, 17, 19, 23) is $r=0.986$.

Rampino and Stothers then go on to compare the gaps in the two columns by Student's t-test for matched pairs. There are two serious difficulties in this; the point will be clearer as we introduce some notation. Let $X_1=11,\ldots,X_9=245$ be the entries in the first column and $Y_1=0,\ldots,Y_9=259$ be the entries in the second column. Then Rampino and Stothers apply Student's t-test to the eight differences $Z_i=(Y_{i+1}-Y_i)-(X_{i+1}-X_i)$. The test statistic is then the average of the Zs divided by what would be an appropriate estimate of standard error of that average if the Zs were independent. But the average of the Zs reduces algebraically to $\bar{Z}=[(Y_9-Y_1)-(X_9-X_1)]/8$; it does not depend on the intermediate values at all. If the two columns spanned the same time interval, it would reduce identically to zero, regardless of the intermediate values. In other words, the Zs are highly correlated because adjacent intervals share common endpoints; the test as performed would be valid only if both series were random walks (hardly a tenable hypothesis, particularly for galactic plane crossings). Consequently, the denominator of the t-statistic grossly over estimates the standard error of the numerator and biases the test statistic towards zero. Even if their test were valid, there is a second difficulty in this aspect of their argument: they interpret a small t-value as evidence in favour of a null hypothesis, when it could as well be held to reflect a paucity of data and a test insufficiently powerful to detect a difference.

S. M. STIGLER

Department of Statistics, University of Chicago, Chicago, Illinois 60637, USA

1. Rampino, M. R. & Stothers, R. B. *Nature* 308, 709–712 (1984).

RAMPINO AND STOTHERS REPLY— Stigler regards as misleading the two statistical tests that we[1] made of the correlation between two observed time series, one containing the dates of the last nine crossings of the galactic plane by the Solar System[2], and the other consisting of the dates of the last nine major mass-extinction episodes on Earth, which we selected from the marine extinctions record published by Raup and Sepkoski[3]. Both records cover the past 260 Myr.

In our first test, we computed the correlation coefficient between the two series and obtained $r=0.996$. Stigler correctly points out that any two monotonically increasing series have a numerically high correlation. As a random test case, he correlates the first nine prime numbers with our mass-extinctions time series and obtains $r=0.986$. To make a better comparison, however, we have computed 5,000 random time series, each containing nine dates drawn from a fixed time range (0–260 Myr). These series have been correlated with our mass-extinctions time series. We find agreement with Stigler in that a statistically significant percentage of cases (13%) have $r \geqslant 0.986$.

On the other hand, only an insignificant 0.4% of the cases computed in these Monte Carlo simulations show $r \geqslant 0.996$. Furthermore, if we correlate our random time series with the galactic time series, the percentage is even lower, 0.2%. Stigler's mistake, therefore, is to regard $r=0.996$ as being close to $r=0.986$ in the case of two monotonically increasing series with nine members each.

With regard to our second test, the matched-pairs t-test applied to the time intervals, Stigler again makes an invalid criticism, because the data for any matched-pairs t-test at all can be recast and analysed in exactly the way he presents, with the same conclusion. Consider, for example, two random samples drawn from independent populations and containing eight elements each: $A_1,\ldots,A_8$ and $B_1,\ldots,B_8$. Enumerate a series of nine Xs and a series of nine Ys by computing $A_i=X_{i+1}-X_i$ with any X_1 and $B_i=Y_{i+1}-Y_i$ with any Y_1. The t-test is now straightforwardly applied to the differences $Z_i=B_i-A_i$. Because the test statistic t is proportional to the average of the Zs, we compute $\bar{Z}$ and find after some manipulation

$$\bar{Z}=[(Y_9-Y_1)-(X_9-X_1)]/8$$

which is identical to Stigler's result. Thus, $\bar{Z}$ appears not to depend at all on the intermediate values of the Xs and Ys, as Stigler noted. Stigler's result, therefore, is seen to be just a mathematical rearrangement of the data. As long as the sample elements, the As and Bs, are independent and approximately normally distributed, the t-test is valid. In our case, the sample elements are the time intervals in the time series, which are the relevant, physically independent units and are affected by random errors arising from many complex physical and accidental factors for both series[1–3]. The time intervals are, in fact, found to be approximately normally distributed. Thus, these time series can be regarded as random walks. From the point of view of testing the equality of the averages of the time intervals in the two series, it makes no difference in what order the time intervals are taken, although, if pairs are to be matched, the members of pairs must be kept together.

Our use of the matched-pairs t-test appears, therefore, to be entirely valid. This test is usually considered to be a powerful one, and one for which a sample size of eight is not unusual[4]. (Student[5] himself used sample sizes of as small as two.) Our original result, based on a two-tailed test, was $t=0.91$ with 7 d.f., therefore $P=0.39$. If, however, we do not pair the time intervals but do allow for the unequal variances in the two samples of time intervals and for the small sample sizes[6], we obtain $t=0.82$ with effectively 6 d.f., and so $P=0.45$. Alternatively, by simply testing the hypothesis that the average of the time intervals in the mass-extinctions series equals 33 Myr (which is the average time interval in the galactic series), we find $t=1.01$ with 7 d.f., so $P=0.35$. Clearly, even large differences in our test assumptions make little difference to the results.

Thus, we cannot reject the null hypothesis that the averages of the time intervals in the two series are equal. On the other hand, the tests cannot tell us how nearly equal the averages may be. A feeling for what can be safely rejected follows from a further consideration: with the same number of time intervals, another mass-extinctions time series might have a considerably different length. To demonstrate the consequences of adopting the original time series given by Raup and Sepkoski[3], in which the last nine mass-extinction episodes occurred at 11.3, 38, 65, 91, 125, 144, 163, 175 and 194 Myr ago, we have made a new correlation with the galactic time series, finding $r = 0.994$ and, for the matched time intervals, $t = 4.94$ with 7 d.f., for which $P < 0.01$. Thus, the two series are strongly correlated, but in this case they have significantly different averages of the time intervals.

Note that the statistical tests in our original paper were not central to the search for periodicities or the physical arguments there. The approximate periodicity that appeared in both of our listed time series was formally detectable because the variance of the distribution of time intervals in each series was sufficiently small.

MICHAEL R. RAMPINO
RICHARD B. STOTHERS
National Aeronautics and Space Administration, Goddard Space Flight Center, Institute for Space Studies, 2880 Broadway, New York, New York 10025, USA

1. Rampino, M. R. & Stothers, R. B. *Nature* **308**, 709–712 (1984).
2. Innanen, K. A., Patrick, A. T. & Duley, W. W. *Astrophys. Space Sci.* **57**, 511–515 (1978).
3. Raup, D. M. & Sepkoski, J. J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 801–805 (1984).
4. Klugh, H. E. *Statistics, the Essentials for Research* Ch. 10 (Wiley, New York, 1970).
5. Student [Gossett, W. S.] *Biometrika* **6**, 1–25 (1908).
6. Bailey, D. E. *Probability and Statistics* Ch. 17 (Wiley, New York, 1971).

Hund's rule

BOYD[1] recently suggested an explanation of Hund's rule in terms of electron correlation. A previous investigation of this approach[2] gave rise to a richer interpretation of two-electron spectra in terms of an alternating rule, in place of the usual Hund rule first noted by Russell and Meggers[3] for the spectrum of scandium. The alternating rule states that, within a given configuration, the energy ordering of singlet and triplet states reverses each time the angular momentum changes by one unit. For the term of greatest angular momentum, the triplet lies lowest. The rule reliably orders singlet–triplet energy levels in ~90% of known cases.

The alternating rule follows from the assumption that the state with the larger average value of the inter-electronic angle, θ_{12}, has less shielding and therefore a lower energy. This is not far from Boyd's approach. One might expect a triplet always to have less shielding, owing to the Fermi hole, but wavefunction antisymmetry can give rise to features in the singlet which resemble the triplet hole[2]. As the charge density $\rho(\theta_{12})$ has been found to obey the alternating rule[2], the Fermi hole argument is too simplistic.

The alternating rule helps in explaining cases deviating from Hund's first rule, but many exceptions are known. When the second rule fails, a singlet can be the lowest term in a configuration. An example is the 2p3d configuration of carbon, where the ^{1}D lies below the ^{3}F, contrary to Hund's first rule. The alternating rule, however, is obeyed: the ^{1}D lies below the ^{3}D, and the ^{3}F is below the ^{1}F. For the npn'd configuration, Hund's rules work only in ~50% of the cases[2] and many of the exceptions may be caused by a failure of Hund's second rule.

Recent studies of electron distributions in two-electron atoms[4,5] are beginning to yield a new and more comprehensive picture. Some states, for example, the doubly-excited states of helium-like atoms, exhibit very strong angular correlations, much greater than in singly-excited states, to the extent that collective, quasi-molecular quantization, with near-rigid rotations and bending and stretching vibrations, describes the electron distribution. Even the alkaline-earth atoms and alkali negative ions show considerable A—B—A molecular behaviour, although these systems have less rigid structures than He**.

That radial correlation sometimes dominates the energy orderings is difficult to reconcile with shielding concepts[2]. The 2s3s states of helium have similar angular electron distributions, much like that of the 2s^2, ^{1}S state, but the radial distributions are all quite different. Both the 2s3s states have strong angular correlation expected for a linear quasi-molecular system, with $\rho(\theta_{12})$ centred around π. The triplet corresponds to one quantum of excitation in the antisymmetric stretching mode for the A—B—A molecule; it has a lower energy than the singlet because the antisymmetric stretching has a lower frequency when A is lighter than B. The 2s2p, ^{3}P helium state, with a strong maximum in $\rho(\theta_{12})$ near π, corresponds to the first-excited rotor state, with no bending excitation. The angular distribution in the 2s2p, ^{1}P state is very different, with the maximum at slightly more than $\pi/2$, and with finite probability density at $\theta_{12} = 0$. This singlet is one of two partners in a nearly degenerate pair with one quantum in the bending vibration; the other is the lower energy 2p^2, ^{3}P, having almost the same angular distribution, but no electron density at $\theta_{12} = 0$.

In the quasi-molecular picture, the appropriate comparisons, as exemplified above, may be between states of different configurations. When electron correlation and configuration interaction are large, it is inappropriate to classify states in terms of a one-electron configuration, and Hund's rule loses its meaning. Nevertheless, it is interesting to ask under what circumstances a shielding argument can predict the energy orderings. The alternating rule has many exceptions for the npn'd, P and D terms, and future considerations of this problem should perhaps focus on physical interactions, such as collective motion kinematics and shielding, rather than on quantum numbers associated with a specific model and representation.

J. W. WARNER
R. STEPHEN BERRY
Department of Chemistry, The University of Chicago, Chicago, Illinois 60637, USA

1. Boyd, R. J. *Nature* **310**, 480–481 (1984).
2. Warner, J. W., Bartell, L. S. & Blinder, S. M. *Int. J. Quantum Chem.* **18**, 921–932 (1980).
3. Russell, N. H. & Meggers, W. F. *Sci. Pap. natn Bur. Stand.* **22**, 364–371 (1927).
4. Rehmus, P. D. & Berry, R. S. *Chem. Phys.* **38**, 247–275 (1979).
5. Ezra, G. & Berry, R.S. *Phys. Rev.* **A28**, 1974–1988 (1983).

Frost rings in trees and volcanic eruptions

LAMARCHE and Hirschboeck[1] find good agreement between the timing of notable frost events in the western United States and the occurrence of major volcanic eruptions worldwide from AD 1500 onwards. Notable frost events are defined as those damaging trees (seen in tree rings) at two or more localities or in 50% or more of sampled trees in any one locality. For the periods 1882–1968 and 1500–1880 they find, by considering coincidences of these phenomena in 3-yr segments, χ^2-values of 16.0 and 14.7, respectively, in frost versus volcano contingency tables, indicating significance at the 99.95% confidence level.

I have the following comments to make: a recomputation of χ^2 is possible using only the occasions (underlined in Table 1 of ref. 1) of geographically-widespread frost damage (that is, in both the Great Basin and the Rocky Mountains regions, >10° longitude apart). The results are: 1881 onwards, $\chi^2 = 6.7$, just significant at 99% confidence level; 1500–1880, $\chi^2 = 7.8$, significant at 99% confidence level.

The authors' selection of volcanic events, after Lamb[2], appears to be inconsistent at a few points. In general, they selected events printed[2] in bold type, indicating a consensus among compilers as to the importance of the events, with DVI/$E_{max} \geq 1{,}000$ for the eruption or combined eruptions. Events, with the latter parameter in bold type and $\geq 1{,}000$ but with the name of the event in normal type (for example, Pogrumni, 1795), were omitted. To have followed this objective procedure consistently would have been acceptable but the following deviations were noted in particular: (1) Awu and Gunung (both 1641, combined DVI/$E_{max} = 1{,}500$) were omitted. (2) The 1785 eruption of Vesuvius was not assigned a DVI by Lamb and should have been omitted. (3) Hekla (1845) and Amargura (1846) (combined DVI/$E_{max} \simeq 1{,}800$) were omitted. (4) A group of four eruptions (1886–88) (combined DVI/E_{max} = 1,100 (estimated)[2] in 1888–90) was omitted. Item (2) will not affect the authors' results because Vesuvius was grouped with other eruptions, themselves with DVI/$E_{max} > 1{,}000$. However, according to Table 1 in ref. 1, the events in (1), (3) and